

AMENDMENTS TO THE CLAIMS

Please amend the claims as follows:

- 1 1. (Currently amended) A computer-implemented method of text equivalenc-
2 ing from a string of characters comprising:
3 modifying the string of characters using a predetermined set of heuristics;
4 [comparing] performing a character-by-character comparison of the modified
5 string with a known string of characters in order to locate a match;
6 responsive to not finding an exact match, forming a plurality of sub-strings of
7 characters from the string of characters; and
8 using an information retrieval technique on the sub-strings of characters to de-
9 termine a known string of characters equivalent to the string of characters.
- 1 2. (Original) The method of claim 1, wherein the information retrieval tech-
2 nique further comprises:
3 weighting the sub-strings;
4 scoring the known string of characters; and
5 retrieving information associated with the known string of characters with the
6 highest score.
- 1 3. (Original) The method of claim 2, further comprising, responsive to the
2 highest score being greater than a first threshold, automatically accepting the known
3 string of characters as an exact match.

1 4. (Original) The method of claim 2, further comprising, responsive to the
2 highest score being less than a second threshold and greater than a first threshold,
3 presenting the known string of characters to a user for manual confirmation.

1 5. (Original) The method of claim 2, further comprising, responsive to the
2 highest score being less than a second threshold and greater than a third threshold,
3 presenting the known string of characters to a user to select the equivalent string of
4 characters.

1 6. (Original) The method of claim 1, wherein the sub-strings of characters are
2 3-grams.

1 7. (Original) The method of claim 1, wherein the string of characters is se-
2 lected from the group consisting of a song title, a song artist, an album name, a book
3 title, an author's name, a book publisher, a genetic sequence, and a computer pro-
4 gram.

1 8. (Original) The method of claim 1, wherein the predetermined set of heuris-
2 tics comprises removing whitespace from the string of characters.

1 9. (Original) The method of claim 1, wherein the predetermined set of heuris-
2 tics comprises removing a portion of the string of characters.

1 10. (Original) The method of claim 1, wherein the predetermined set of heu-
2 ristics comprises replacing a symbol in the string of characters with an alternate rep-
3 resentation for the symbol.

1 11. (Original) The method of claim 1 further comprising storing an indication
2 that the string of characters is the equivalent of the known string of characters.

1 12. (Currently amended) A computer implemented system for text equiva-
2 lencing from a string of characters comprising:

3 a heuristics module for modifying the string of characters using a predeter-
4 mined set of heuristics;

5 a comparator module, coupled to the heuristics module, for [comparing] per-
6 forming a character-by-character comparison of the modified string with a
7 known string of characters in order to find a match;

8 a sub-string formation module, coupled to the comparator module, responsive
9 to not finding an exact match, for forming a plurality of sub-strings of
10 characters from the string of characters; and

11 an information retrieval module, coupled to the sub-string formation module,
12 for performing an information retrieval technique on the sub-strings of
13 characters to determine a known string of characters equivalent to the
14 string of characters.

1 13. (Original) The system of claim 12, wherein the information retrieval
2 module further comprises:
3 a weight module for weighting the sub-strings;
4 a score module for scoring the known string of characters; and
5 a retrieval module, coupled to the weight and score modules, for retrieving in-
6 formation associated with the known string of characters with the highest
7 score.

1 14. (Original) The system of claim 13, further comprising an accept module,
2 coupled to the retrieval module, for accepting the information retrieved as an exact
3 match for the highest score greater than a first threshold.

1 15. (Original) The system of claim 13, further comprising an accept module,
2 coupled to the retrieval module, for presenting the information retrieved to a user for
3 manual confirmation for the highest score less than a first threshold and greater than
4 a second threshold.

1 16. (Original) The system of claim 13, further comprising an accept module,
2 coupled to the retrieval module, for presenting the information retrieved to the user
3 as a set of options for a user to select for the highest score less than a second thresh-
4 old and greater than a third threshold.

1 17. (Original) The system of claim 12, wherein the sub-strings of characters
2 are 3-grams.

1 18. (Original) The system of claim 12, wherein the string of characters is se-
2 lected from the group consisting of a song title, a song artist, an album name, a book
3 title, and author's name, a book publisher, a genetic sequence, and a computer pro-
4 gram.

1 19. (Original) The system of claim 12, wherein the predetermined set of heu-
2 ristics comprises removing whitespace from the string of characters.

1 20. (Original) The system of claim 12, wherein the heuristics module com-
2 prises a removal module for removing a portion of the string of characters.

1 21. (Original) The system of claim 12, wherein the heuristics module com-
2 prises a replacement module for replacing a symbol in the string of characters with
3 an alternate representation for the symbol.

1 22. (Original) The system of claim 12 further comprising a database update
2 module for storing an indication that the known string of characters is the equivalent
3 of the known string of characters.

1 23. (Currently amended) A computer-readable medium comprising com-
2 puter-readable code for performing text equivalencing from a string of characters
3 comprising:

4 computer-readable code adapted to modify the string of characters using a
5 predetermined set of heuristics;

6 computer-readable code adapted to [compare] perform a character-by-
7 character comparison of the modified string with a known string of charac-
8 ters in order to locate a match;

9 computer-readable code, responsive to not finding an exact match, adapted to
10 form a plurality sub-strings of characters from the string of characters; and
11 computer-readable code adapted to use an information retrieval technique on
12 the sub-strings of characters to determine a known string of characters
13 equivalent to the string of characters.

1 24. (Original) The computer-readable medium of claim 23, wherein the in-
2 formation retrieval technique further comprises:

3 computer-readable code adapted to weight the sub-strings;

4 computer-readable code adapted to score the known string of characters; and

5 computer-readable code adapted to retrieve information associated with the
6 known string of characters with the highest score.

1 25. (Original) The computer-readable medium of claim 24, further compris-
2 ing computer-readable code, responsive to the highest score being greater than a first
3 threshold, adapted to automatically accept the known string of characters as an exact
4 match.

1 26. (Original) The computer-readable medium of claim 24, further compris-
2 ing computer-readable core, responsive the highest score being less than a second
3 threshold and greater than a first threshold, adapted to present the known string of
4 characters to a user for manual confirmation.

1 27. (Original) The computer-readable medium of claim 24, further compris-
2 ing computer-readable code, responsive to the highest score being less than a second
3 threshold and greater than a third threshold, adapted to present the known string of
4 characters to a user to select the equivalent string of characters.

1 28. (Original) The computer-readable medium of claim 23, wherein the sub-
2 strings of characters are 3-grams.

1 29. (Original) The computer-readable medium of claim 23, wherein the string
2 of characters selected from a group consisting of a song title, a song artist, an album
3 name, a book title, an author's name, a book publisher, a genetic sequence, and a
4 computer program.

1 30. (Original) The computer-readable medium of claim 23, wherein the pre-
2 determined set of heuristics comprises removing whitespace from the string of char-
3 acters.

1 31. (Original) The computer-readable medium of claim 23, wherein the pre-
2 determined set of heuristics comprises removing a portion of the string of characters.

1 32. (Original) The method of claim 23, wherein the predetermined set of heu-
2 ristics comprises replacing a symbol in the string of characters with an alternate rep-
3 resentation for the symbol.

1 33. (Original) The computer-readable medium of claim 23 further comprising
2 updating the known string of characters to indicate the string of characters is the
3 equivalent of the known string of characters.

1 34. (Currently amended) A computer-implemented system for performing
2 text equivalencing from a string of characters comprising:
3 a modifying means for modifying the string of characters using a predeter-
4 mined set of heuristics;
5 a comparator means for [comparing] performing a character-by-character
6 comparison of the modified string with a known string of characters in or-
7 der to locate a match;

8 responsive to not finding an exact match, a formation means for forming a
9 plurality sub-strings of characters from the string of characters; and
10 an information retrieval means for determining a known string of characters
11 equivalent to the string of characters.

1 35. (Original) The system of claim 34, wherein the information retrieval
2 means further comprises:
3 a weight means for weighting the sub-strings;
4 a score means for scoring the known string of characters; and
5 a retrieval means for retrieving information associated with the known string
6 of characters with the highest score.